# NSA Surveillance: Exploring the Geographies of Internet Interception

Andrew Clement[1]

[1] Faculty of Information, University of Toronto

**Abstract**

The National Security Agency's various surveillance programs recently revealed by Edward Snowden are collectively arguably the largest personal data collection and analysis operation in history. While the foremost exemplar of a fine-grained, global information system, they also represent among the most serious contemporary challenges to democratic governance and civil liberties. Based on media coverage and leaked secret documents, this paper analyses the main NSA data interception programs and their geographic characteristics. This research also draws on IXmaps.ca, a crowd-sourced, interactive mapping application to show internet users where their personal traffic may be intercepted by the NSA. In particular, it demonstrates that internet surveillance facilities located in relatively few strategic locations enable a nearly comprehensive collection of domestic U.S. internet traffic.

**Keywords:** NSA surveillance, warrantless wiretapping, internet surveillance, privacy

**Citation**: Clement, A. (2014). NSA Surveillance: Exploring the Geographies of Internet Interception. In *iConference 2014 Proceedings* (p. 412–425). doi:10.9776/14119



**Acknowledgements**: The IXmaps project is the work of a research team that currently includes Jonathan Obar, Colin McCann and Antonio Gamba. David Phillips, Steve Harvey, Gabby Resch, Erik Stewart, Nancy Paterson, Misha Snyder and Lauren DiMonte have made invaluable contributions at earlier stages of the project. We are also grateful to those individuals, largely anonymous, who have contributed to the database by installing and running TRgen, or have provided feedback that has helped improve the application. This research has received funding from Canada's Social Sciences and Humanities Research Council and the Office of the Privacy Commissioner of Canada.

**Contact**: andrew.clement@utoronto.ca

## 1 Introduction

The 2013 revelations of U.S. National Security Agency (NSA) surveillance programs brought to public attention by whistleblower Edward Snowden's release of hitherto secret internal documents have sparked a storm of controversy. Their breathtaking scope, scale, and questionable legality, largely confirm the earlier allegations of clandestine domestic spying by retired AT&T technician Mark Klein(2009), author James Bamford (2008), and others (Landau, 2011). Klein, reported in 2006 that the NSA had secretly installed surveillance equipment in AT&T's main San Francisco internet exchange point capable of copying and analyzing potentially all the internet traffic passing through that location. He indicated that similar facilities were operating in other AT&T switching centres. Termed NSA 'warrantless wiretapping', this was arguably the largest single state surveillance program conducted over the communications of its citizens (Bamford, 2008). Congressional passage of special legislation in 2008 to protect telecommunications carriers against the dozens of lawsuits that ensued largely ended media attention to the controversy at that time, but left the constitutional and civil liberties issues unresolved, and considerable mystery remaining about what the NSA is actually doing.

Thanks to Snowden, as well as the reporters he handed the trove of 58,000 NSA secret documents to, it is now much clearer that the warrantless wiretapping program was the tip of a much bigger iceberg, covering all forms of telecommunications traffic and implicating most of the major telecommunications carriers across the U.S. For the first time in decades, the widely publicized revelations have prompted a lively discussion in the U.S. as well as elsewhere about the appropriateness and legality of NSA surveillance. While the almost weekly breaking news stories are producing an increasingly detailed collage of hitherto secret NSA operations, there is still much more we need to learn and understand to have the informed

public debate now long overdue. Part of the challenge is to make sense of the welter of fragmentary reports, to yield a more coherent picture of the NSA's vast surveillance infrastructure. This paper seeks to contribute to the public debate now getting underway by shedding light on the widespread state surveillance of everyday citizen internet communication. It focusses especially but not exclusively on internet surveillance within the U.S., and not so much on the vital legal, political and moral issues, but more on the technical, geographically specific aspects – what data is collected, whose personal data, where and how.

During the Cold War, the NSA, as primarily a signal intelligence operation, concentrated on capturing over-the-air transmissions, such as via satellite or micro-wave relay, that could be intercepted passively by setting up antennae within broadcast range. Tapping into wireline communication was similar in that signals in analog transmissions typically 'leaked' and sensitive receivers placed along the lines could pick up the transmissions. The widespread shift during the 1990s to digital networks, notably using fibre optic cables, rendered the conventional interception modes obsolete and there were concerns within intelligence agencies about "going dark."(Bamford, 2008) Because there is typically little or no electro-magnetic or optical signal leakage from digital transmissions, interception initially became much more challenging, not only technically, but also politically and organizationally. This is because it meant physically gaining access to the transmission equipment or breaking into the transmission path. However, just as the capabilities of digital communications, storage and analysis have soared, so too has the capacity to surveil digital activity on a massive scale. Besides the extraordinary technical prowess the U.S. is able to deploy in the service of its perceived surveillance and security needs, the U.S. also has a strategic geographic advantage in that a disproportionate share of international data communications – an advantage the NSA is well aware of. See Figure 1.

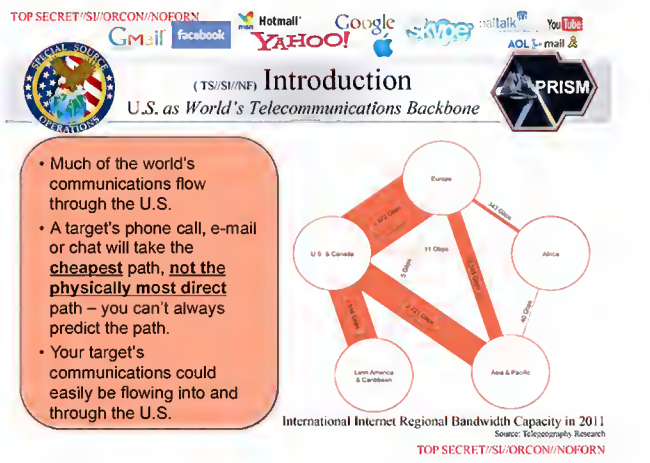


Figure 1: U.S. as World's Telecommunications Backbone[1]

These expansive changes in surveillance technique have been driven not only by advances in the underlying information technologies, but also by geopolitical changes in the nature of threats perceived by national security and law enforcement agencies and what are appropriate means to address them. These changes also have important implications for how (potentially) surveilled subjects can respond to them and more generally how democracies can govern the powerful forces unleashed by them. In particular, a geographic perspective on NSA surveillance, and its interception capabilities in particular, is helpful in understanding where it can be done, what parties are implicated, what legal jurisdictions apply and how relations among the various diverse actors distributed across the nation and around the world are affected.

[1] Source: *Washington Post*, NSA slides explain the PRISM data-collection program
June 6, 2013, Updated July 10, 2013. http://www.washingtonpost.com/wp-srv/special/politics/prism-collection-documents/

This paper, drawing on surveillence studies perspectives (Lyon, 2007), views NSA interception not as an isolated occurance, but as reflecting a wider societal phenomenon, in which surveillance, "monitoring people in order to regulate or govern their behaviour" (Gilliom & Monahan, 2013, p.2) as a central organizing principle. Surveillance is often benign, even essential, but is becoming so pervasive and inextricably connected to everyday activities that we can characterise our contemporary 'western' life as a surveillance society. At the same time, it is important to recognize that notwithstanding its burgeoning extent and intensity, surveillance and its effects are not uniform, affecting everybody, everywhere in the similar ways. In these respects, the NSA surveillance programs offer an albeit extreme, potentially malignant but nevertheless revealing exemplar of systemic surveillance trends in our increasingly digitally mediated contemporary society.

We begin by describing briefly the NSA's most prominent programs, with a focus on those that collect the raw data on which subsequent analysis and action is based. This provides the opportunity to highlight the specific kinds of personal data that are collected and from whom, where. The evidence for these programs comes almost exclusively from secret NSA documents released by Snowden, as reported from June 5 to the end of December 2013, mainly in the Guardian and Washington Post newspapers. The authenticity of these leaked documents, with very few exceptions, is generally acknowledged, including by the US government. The following section examines in more detail the NSA internet surveillance conducted in the US at major internet exchange points, referred to as the 'warrantless wiretapping' program. We again rely mainly on mainstream news reports, but in the period 2005-2012. The third section builds on what was reported at that time about the location and nature of the NSA US domestic internet surveillance facilities to explore empirically whether individually generated internet routes may be exposed to NSA warrantless wiretapping. This is done using a research-based internet mapping tool known as IXmaps, developed to map internet exchange points and the traffic routed through them. The software tool found at IXmaps.ca[2] aggregates crowd-sourced internet users' 'traceroutes' and shows them where their personal traffic is likely to have been intercepted by the NSA. In contradistinction to the common metaphor of the internet as a space-less, featureless 'cloud', we demonstrate that with interception points in relatively few major cities (<20) the NSA is capable of intercepting a large proportion (>95%) of domestic internet traffic. We close by reflecting on the role that a geographic visualization tool may play in facilitating public understanding of internet surveillance and by calling for discussion within the information studies field about its 'darker side'.

## 2 NSA Surveillance Programs

The reporting of NSA surveillance activities, beginning in June 2013 when whistleblower Edward Snowden handed over secret documents to reporters Glenn Greenwald and Laura Poitras, for the first time brought to wide public attention details of the NSA's comprehensive array of data collection, archiving, mining, analysis and visualization programs. While much of the legal and political controversy these revelations provoked has focused so far, at least in the U.S., on whether access to collected data about Americans is legal or even constitutional, of wider significance is the existence of a global apparatus designed for and capable of intercepting virtually all electronic communications, ostensibly for 'security' purposes. These data accumulation activities in aggregate are enormous. To give an idea of their scope, William Binney, a former NSA mathematician and Technical Director of the World Geopolitical and Military Analysis Reporting Group, estimated in 2012 that the agency had "assembled on the order of 20tn transactions about US citizens with other US citizens", and this included "only ... phone calls and emails". In 2010, well before the Snowden revelations, according to this same *Guardian* article, the *Washington Post* reported that

[2] http://ixmaps.ca

> "every day, collection systems at the [NSA] intercept and store 1.7bn emails, phone calls and other type of communications."[3]

More recently, this paper reported that the NSA "is gathering nearly 5 billion records a day on the whereabouts of cellphones around the world."[4] The breathtaking, global geographic scope of the NSA's data gathering capabilities is demonstrated in its map of the 'Worldwide SIGINT/Defense Cryptologic Platform (See Figure 2).

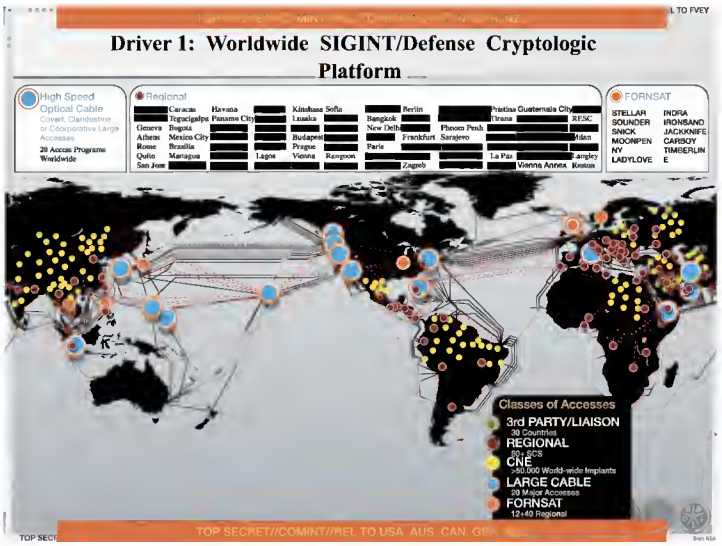


Figure 2: NSA's worldwide data gathering surveillance infrastructure [5]

Whose data the NSA can capture, where and how has now become a topic of widespread concern. While questions of what is actually done with the personal data the NSA collects, and what protections may or may not exist, at least for "US persons", is of course at least as important, we focus here on the fundamental issue of original capture technique and locale. Of the numerous and varied NSA programs we so far now know something about, we'll explore here five that reveal most directly the geographic scope and personal informational details of mass NSA surveillance. We begin with two analysis and visualization programs, *Boundless Informant* and *X-Keyscore*, that give a broad overview of the scale and intensity of surveillance, and then examine more closely three programs directly related to interception, each of which adopt a distinctive technique: bulk telephony meta-data collection, *Prism*, and *Upstream*.

## 2.1 Boundless Informant

The aptly named, top secret, Boundless Informant program is a data mining tool described in an official Global Access Operations (GAO) FAQ, as providing "the ability to dynamically describe GAO's collection capabilities (through metadata record counts) with no human intervention and graphically display the information in map view, bar chart or simple table".[6] A GAO slide presentation claims it uses 'Big Data

[3] Glenn Greenwald, XKeyscore: NSA tool collects 'nearly everything a user does on the internet', Guardian, 31 July 2013. http://www.theguardian.com/world/2013/jul/31/nsa-top-secret-program-online-data

[4] Barton Gellman and Ashkan Soltani, NSA tracking cellphone locations worldwide, Snowden documents show, Washington Post, 4 December, 2013.http://www.washingtonpost.com/world/national-security/nsa-tracking-cellphone-locations-worldwide-snowden-documents-show/2013/12/04/5492873a-5cf2-11e3-bc56-c6ca94801fac_story.html?hpid=z1

[5] Floor Boon, Steven Derix and Huib Modderkolk, NSA infected 50,000 computer networks with malicious software, 23 november 2013, NRC.NL, http://www.nrc.nl/nieuws/2013/11/23/nsa-infected-50000-computer-networks-with-malicious-software/ Note that this image does not show the domestic interception within the Five Eye countries, which we'll discuss later in relation to the U.S.

[6] Guardian, Boundless Informant: NSA explainer – full document text, June 8, 2013
http://www.theguardian.com/world/interactive/2013/jun/08/boundless-informant-nsa-full-text

technology to query SIGINT [signal intelligence] collection in the cloud to produce near real-time intelligence describing the agency's available SIGINT infrastructure and coverage."[7] While this tool doesn't access all the information the NSA collects (e.g. information covered by FISA restrictions is not included), the volume is impressive. The *Guardian* reports that "in March 2013 the agency collected 97bn pieces of intelligence from computer networks worldwide", including nearly 3bn in the U.S. See Figure 3 for a partial 'heat map' showing relative amounts of data collected in various countries that month. While a few countries show particularly intense collection (marked in red), this map shows that interception is widespread around the globe.

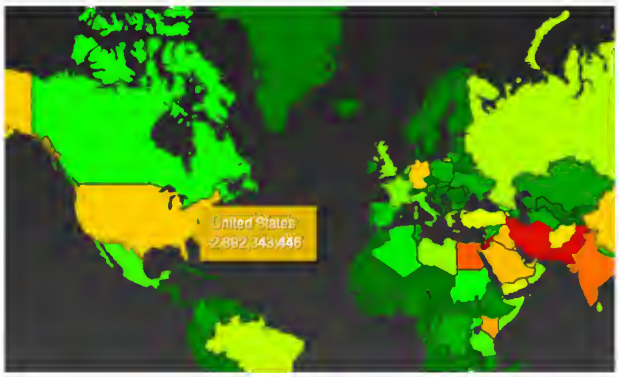


Figure 3: Heat map of NSA meta data collection in March 2013[8]

## 2.2 X-KeyScore

X-KeyScore is a query tool designed to allow authorized analysts to interrogate through a desktop interface the NSA's vast world-wide intelligence holdings. 'Selectors', such as an email address or IP address, can be used to access stored data as well as initiate "ongoing 'real-time' interception of an individual's internet activity." Due the large volumes collected, data is initially held close to the point of capture and much of it is deleted after a few days.[9] Figure 4 shows the world-wide distribution of these caches, again showing the broad geographic scope of NSA surveillance. It is not just NSA analysts who have access to this data, but as the top of the slide shows, the security agencies in the other "Five Eyes" countries – Australia, Canada, Great Britain and New Zealand. There are also reports of access by intelligence services in other US allies such as Germany[10] and Isreal.[11]

[7] Guardian, Boundless Informant NSA data-mining tool – four key slides, June 8, 2013. http://www.theguardian.com/world/interactive/2013/jun/08/nsa-boundless-informant-data-mining-slides

[8] Source: Glenn Greenwald and Ewen MacAskill, Boundless Informant: the NSA's secret tool to track global surveillance data, *Guardian*, 11 June 2013. http://www.theguardian.com/world/2013/jun/08/nsa-boundless-informant-global-datamining

[9] Sean Gallagher, Building a panopticon: The evolution of the NSA's XKeyscore, *ArsTechnica*, Aug 9 2013. The need for these globally distributed caches may well be temporary, given the massive Utah data centre that Bamford reported on in 2012, originally scheduled to open September 2013. See: Bamford (2012).

[10] *Der Spiegel.* "'Prolific Partner': German Intelligence Used NSA Spy Program", July 20, 2013.

[11] Glenn Greenwald, Laura Poitras and Ewen MacAskill, NSA shares raw intelligence including Americans' data with Israel. *Guardian,* 11 September 2013. http://www.theguardian.com/world/2013/sep/11/nsa-americans-personal-data-israel-documents


Figure 4: Location of data caches accessible by X-Keyscore[12]

We turn now to looking at the major surveillance programs revealed so far that generate these vast troves of intelligence data.

## 2.3 Bulk telephony meta-data collection

The NSA has long been suspected of collecting the telephone calling records of Americans, and there have been court cases challenging this practice dating back to 2008. The main charges were against AT&T as well as Verizon/MCI, BellSouth, Sprint, and Cingular.[13] However, the FISA Amendments Act (FISAA) of 2008, popularly known as the "Telecom Immunity Act", rendered these cases moot, as this legislation "allow[s] federal judges to waive lawsuits if the telecom firms can prove that they were authorized by the president and assured that the program was legal."[14] There were several more court cases against the federal government, with the federal government seeking so far successfully to have each dismissed on 'national security' grounds, or because plaintiffs couldn't establish 'standing', i.e. they weren't able to prove that their telephone activities had been intercepted because the existence (or not) of such surveillance was itself a secret.[15] However, the release, first by the *Guardian* and subsequently by the federal government, of the FISA Court order requiring Verizon to "produce to the [NSA] ... on a daily basis ... an electronic copy of ... all call detail records or 'telephony metadata'",[16] confirmed the existence of the secret program and re-ignited the court challenges.[17] Not just international calls are included, but all local calls as well. In other

[12] Glenn Greenwald, XKeyscore: NSA tool collects 'nearly everything a user does on the internet', *Guardian*, 31 July 2013. http://www.theguardian.com/world/2013/jul/31/nsa-top-secret-program-online-data

[13] Unlike the five telecom carriers that faced lawsuits, Qwest reportedly did not comply with the NSA's request to turn in customers' telephone records. Moreover, Qwest CEO claims that this request was made in February 2001, well before the attacks of September 11. In addition to requests for phone records, Qwest was also approached by unnamed "clandestine agencies" about allowing the latter the use of Qwest's "fiber-optic communications network for government purposes." Qwest says that it did not comply.

[14] M. Soraghan, "House passes FISA overhaul" The Hill,June 20, 2008; and
http://www.sourcewatch.org/index.php?title=FISA_Amendments_Act_of_2008

[15] The two most prominent of these cases are *Jewel v. NSA* and *Clapper v. Amnesty*. In *Jewel v. NSA*, EFF is suing the NSA and other government agencies on behalf of AT&T customers to stop the illegal, unconstitutional and ongoing dragnet surveillance of their communications and communications records. In *Clapper v. Amnesty et al* the Supreme Court in 2013 denied the ACLU's challenge to the constitutionality of FISAA based on 'lack of standing.'

[16] *Guardian*, Verizon forced to hand over telephone data – full court ruling, June 6, 2013.
http://www.theguardian.com/world/interactive/2013/jun/06/verizon-telephone-data-court-order

[17] The ACLU, a subscriber of Verizon and hence more confident that the issue of standing would not be problem it was in *Clapper v. Amnesty*, returned to court on June 11, 2013, "challenging the constitutionality of the National Security Agency's mass collection of Americans' phone records." https://www.aclu.org/national-security/aclu-v-clapper-challenge-nsa-mass-phone-call-tracking. On December 2013, Judge William Pauley, of the Federal Southern District of New York, rejected their claim in finding this bulk collection did not violate the U.S. Constitution. See: Dan Roberts NSA mass collection of phone data is legal, federal judge rules. *Guardian*, 27 December 2013. http://www.theguardian.com/world/2013/dec/27/judge-rules-nsa-phone-data-collection-legal. However, a week earlier, Judge Richard J. Leon of the Federal District Court for the District of Columbia, had ruled in a similar case that "that the National Security Agency program that is systematically keeping records of all Americans' phone calls most likely violates the

words, details of every call originated or terminated in the U.S. are captured. The granularity of collection is also impressive. According to the secret FISA court order:

> "Telephony metadata includes comprehensive communications routing information, including, but not limited to session identifying information (e.g. originating and terminating telephone number, International Mobile Subscriber Identity (IMSI) number, International Mobile station Equipment Identity (IMEI) number, etc.), trunk identifier, telephone calling card numbers and duration of call." (Vinsen, 2013, p. 2)

Location data can also be included. While metadata does not enjoy the same degree of legal protection as message content, many have argued it can often be just as sensitive and intrusive.[18] Furthermore, persons affected by the court order are 'gagged', i.e. prohibited from telling any unauthorized person about the order. While so far the only order made public is that of Verizon, there is good reason to believe that other similarly large telecom providers, such as AT&T, Sprint and Cingular, have also been served with equivalent orders. As such, we can conclude that without prior suspicion nearly everyone in the U.S. has all their call details routinely reported every day to the NSA.

### 2.4 Prism

Prism, the most recent of the NSA's large scale domestic data collection program, involves "tapping directly into the central servers of nine leading U.S. Internet companies."[19] Starting in 2007, the NSA has arranged for automated access to the servers of Microsoft, Yahoo, Google, Facebook, PalTalk, AOL, Skype, YouTube and Apple. The *Washington Post* also reports that the

> "NSA collects, identifies, sorts and stores at least 11 different types of electronic communications [including] Chats, E-mail, File transfers, Internet telephone [VoIP], Login/ID, Metadata, Photos, Social networking, Stored data, Video, Video conferencing"[20]

Prism appears to have arisen in response both to legal and political pressures when the 'warrantless wiretapping' program came to light, as well as to get around the increasing use of encryption that rendered analysis of message content captured by Upstream more difficult. Because the on-line services of these nine companies are popular globally, and are covered by US law, their users world wide can expect their personal data to be open to inspection by the NSA, with no expectation of legal protection, if outside the U.S.

### 2.5 Upstream

When the recent round of NSA surveillance revelations broke in June 2013, it was the bulk telephony metadata collection and the Prism program, discussed above, that garnered the greatest media attention. However, because of its potentially even wider reach than either of the other two, it is the NSA's Upstream program, revealed later and incidentally, that is arguably the most significant and politically challenging of the Agency's massive data collection programs. There are few Snowden documents yet (as of December 2013) giving specific details, but it is mentioned in an early *Guardian* article focused on the Prism program.

---

Constitution". Charlie Savage, Judge Questions Legality of N.S.A. Phone Records, *New York Times*, 16 December 2013. http://www.nytimes.com/2013/12/17/us/politics/federal-judge-rules-against-nsa-phone-data-program.html?nl=us&emc=edit_cn_20131216&_r=1& Given this disparity in findings and the significance of the issues at stake, the constitutionality of the NSA's bulk telephony metadata collection program will likely go before the Supreme Court.

[18] Jane Mayer, What's the Matter with Metadata?, *New Yorker*, June 6, 2013.
http://www.newyorker.com/online/blogs/newsdesk/2013/06/verizon-nsa-metadata-surveillance-problem.html

[19] Barton Gellman and Laura Poitras, U.S., British intelligence mining data from nine U.S. Internet companies in broad secret progrram, *Washington Post*, June 6, 2013 http://www.washingtonpost.com/investigations/us-intelligence-mining-data-from-nine-us-internet-companies-in-broad-secret-program/2013/06/06/3a0c0da8-cebf-11e2-8845-d970ccb04497_story.html

[20] Barton Gellman and Todd Lindeman, Inner workings of a top-secret spy program, *Washington Post*, June 29, 2013 http://apps.washingtonpost.com/g/page/national/inner-workings-of-a-top-secret-spy-program/282/

A top secret training slide (see Figure 5), with a world map showing submarine traffic patterns as background, summarizes Upstream as "Collection of communications on fiber cables and infrastructure as data flows past."[21] As this quote suggests, there appear to be two main techniques for accessing data networks – installing fiber optic 'splitters' within major internet switches (infrastructure), and where the switch operators are not sufficiently cooperative, taking the technically more challenging route of tapping into the cables at some point along the route between the switches. Since much of the international internet traffic travels by submarine fiber optic cable, this means installing taps at landing stations or even mid ocean.[22] In both forms of interception, deep packet inspection (DPI) is used to examine and store all aspects of the traffic, including meta-data (e.g. to- and from- IP addresses in the packet headers) as well as communicative content (i.e. the packet 'payload'). Since messages, such as for email, are broken into a series of packets for transmission, these need to be reassembled before the actual message content can be analysed.


Figure 5: NSA training slide for Prism program[23]

We turn now to a more detailed discussion of Upstream and the NSA's domestic U.S. 'warrantless wiretapping' program and where it is most likely to have installed its splitter operations.

## 3 Warrantless Wiretapping Program in the U.S.

The *New York Times* first reported the interception of US domestic communications by the NSA in late 2005.[24] But it wasn't until Mark Klein, a recently retired AT&T technician, revealed the existence of a secret 'splitter' operation at 611 Folsom St in San Francisco that the scope and technical details of NSA surveillance came to public light. Klein reported that AT&T had spliced fiber-optic splitters into 16 'peering links' that connected its network with other major carriers and internet exchange points, directing an exact copy of all the traffic passing through these links into a 'secret room' on the 6th floor, Room 641A. Here a

[21] The British communications intelligence agency Government Communications Headquarters (GCHQ) conducts a similar fibreoptic cable interception program by the name of Project Tempora. See: Ewen MacAskill, Julian Borger, Nick Hopkins, Nick Davies and James Ball, "Mastering the internet: how GCHQ set out to spy on the world wide web" *Guardian*, 21 June 2013 http://www.theguardian.com/uk/2013/jun/21/gchq-mastering-the-internet

[22] The nuclear submarine, Jimmy Carter, has been specially modified to conduct these under water cable tapping operations. See: Associated Press, New Nuclear Sub Is Said to Have Special Eavesdropping Ability, *New York Times*, February 20, 2005. http://www.nytimes.com/2005/02/20/politics/20submarine.html?_r=0

[23] James Ball, NSA's Prism surveillance program: how it works and what it can do, *Guardian*, 8 June 2013 http://www.theguardian.com/world/2013/jun/08/nsa-prism-server-collection-facebook-google

[24] J. Risen and E. Lichtblau, Bush Lets U.S. Spy on Callers Without Courts, *New York Times*, December 16, 2005. http://www.nytimes.com/2005/12/16/politics/16program.html?ex=1145419200&en=87817a067833b164&ei=5070

Narus STA 6400 analyzed all the packets passing by, providing "complete visibility for all Internet applications" according to its vendor. In other words, this operation enables the NSA to monitor not only who is communicating with whom, but potentially the entire contents of these communications as well.

Klein's revelations provoked strong reaction by civil liberties organizations, resulting in over four dozen court cases against U.S. telecom carriers and the federal government. These cases allege that the carriers illegally complied with multiple surveillance requests from the NSA during the Bush Administration to provide without warrants specific information about US citizens.[25]

The secrecy that pervades this topic makes it difficult to determine whether the NSA surveillance program is continuing or not, but the recent reports strongly suggest that not only is it on-going, but is expanding during the Obama Administration. James Bamford's article in the March 2012 issue of Wired details the construction of an enormous data centre in Bluffdale Utah capable of storing and analyzing the complete record of interpersonal internet traffic (Bamford, 2012). In July 2012, three whistleblowers, William E. Binney, Thomas A. Drake, and J. Kirk Wiebe, all former NSA employees, gave evidence in the Electronic Frontier Foundation's (EFF's) (2012) lawsuit against the government's mass surveillance program, Jewel v. NSA in support of the surveillance allegations. In particular, Binney, a former NSA technical director, claims the then current program, known as Stellar Wind, is capable of intercepting virtually all email in the US and much else.[26] The more recent revelations by whistleblower Snowden further confirm the earlier claims, demonstrate that they are part of a much wider suite of surveillance programs and better establish state surveillance as a vital topic for (inter)national debate.

Given that the NSA's internet surveillance is on-going but its details still a closely guarded secret, how can we determine where it is being conducted, and whose traffic is capable of being intercepted? These are the central questions we now examine. We will focus our investigation on AT&T, and the splitter installation at 611 Folsom Street, as this is the best documented case and provides a model for the interception of internet traffic at other major internet exchange points in the U.S. and presumably by other major carriers.

## 4 Mapping NSA surveillance sites and internet traffic through them

### 4.1 Where are the NSA splitter sites?

*While* we know of the NSA splitter site at 611 Folsom Street, what about additional suspected sites? Based on his conversations and meetings with other AT&T technical staff, Klein (2009) reported that similar installations were installed in five other locations – Seattle, San Jose, Los Angeles, San Diego and Atlanta. However, these 6 sites would not be sufficient to comprehensively intercept US internet traffic, as there are other, more important routing centres that would be much more attractive for interception purposes. Scott Marcus, a former Federal Communications Commission expert, estimates that AT&T had 15-20 splitter sites.[27] However, he wasn't able to identify any sites in particular without further specific evidence. Presuming that the NSA's goal was to be able to intercept the largest proportion of US internet traffic with the fewest possible sites (a hypothesis well confirmed by the subsequent Snowden revelations), we developed a crude schema for scoring cities based on how much internet traffic was likely to pass through them. Using only our personal estimates of 3 determinants of internet prominence, with crude relative weightings: telecom infrastructure (10); city size (population) (5); and geographic location in relation to other major

[25] While the Bush Administration initially denied the role of telecommunications carriers, it subsequently confirmed this in general terms. Eric Lichtblau, "Role of Telecom Firms in Wiretaps Is Confirmed", *New York Times,* August 24, 2007. http://www.nytimes.com/2007/08/24/washington/24nsa.html?ex=1345608000&en=4e8428cf3d46306c&ei=5090&partner=rssuserland&emc=rss

[26] P. Harris, US data whistleblower: 'It's a violation of everybody's constitutional rights', *Guardian,* Sept. 15, 2013, http://www.guardian.co.uk/technology/2012/sep/15/data-whistleblower-constitutional-rights

[27] PBS Frontline. Spying on the Home Front, May 15, 2007. http://www.pbs.org/wgbh/pages/frontline/homefront/view/

population centres and telecommunications traffic patterns (4), we developed an ordered ranking of the US cities most likely to host an NSA splitter installation. To test our hypothesis, and more generally provide a means for internet users to see where their data traveled and possibly subject to surveillance, we developed the IXmaps software system. Using a crowd-sourced approach, we invite geographically scattered user to install a customized version of the common traceroute[28] program that populates our database. We add location data for the routers encountered using a variety of standard geo-location techniques and from this users can then selectively map their own or others' traceroutes via a Google Maps mashup. Currently the database contains over 26,000 traceroutes contributed by more than 200 submitters from over 180 originating addresses in North America to in excess of 2600 destination URLs. We examined all the US-only route in the IXmaps database, which currently numbers 2927. Of these, 2839 passed through at least one of the 18 cities we identified as the most likely sites for NSA splitter operations. In other words, installing splitters in the major internet exchange points in just these cities would be sufficient for the NSA to intercept 97% of our US only traceroutes! These are shown in Figure 6.

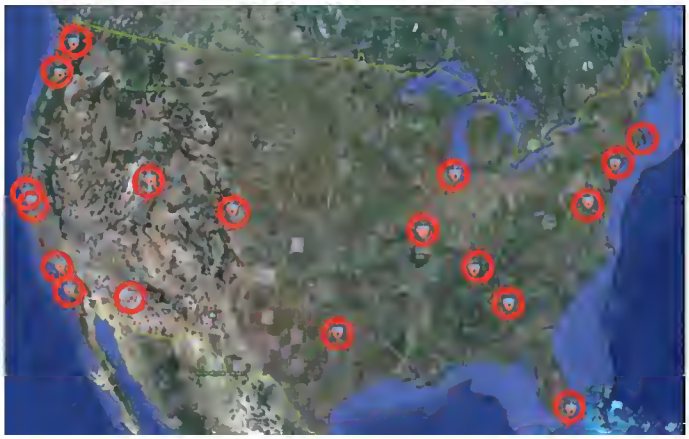

Figure 6: 18 US cities most likely to host NSA splitters[29]

While this result of course does not prove that these cities actually have NSA splitter operations, nor that the NSA has access to all the internet exchange points in them, it is powerful confirmation that if the NSA install splitters in relatively few strategic internet choke points it would be it is technically feasible for it intercept a very large proportion of U.S. internet traffic. This high percentage helps justify our claim that these cities are strongly suspected of hosting NSA warrantless surveillance facilities. It also vividly challenges the popular image of the internet as a 'cloud.'

## 4.2 Does my personal internet data pass through NSA splitter sites?

With the suspected NSA cities identified, we are in the position to give individual internet users a reasoned *indication* of whether their particular communications are likely to be subject to warrantless interception. Exploiting the feature of IXmaps to target any user-provided URL, individuals can generate traceroutes customized to their own internet activities. IXmaps renders both the tabular and map views of these

[28] See: http://en.wikipedia.org/wiki/Traceroute

[29] Biases in the sample of traceroutes contributed by users to the database mean that this particular list of cities and the relative amount of domestic U.S. traffic that could by intercepted by NSA splitters installed in them needs to be treated with caution. The chronic difficulties, widely recognized in the internet routing research community, in accurately geo-locating routers based on hostnames, IP addresses and latencies, further complicate the picture. Nevertheless, we believe the overall conclusions about a relatively small number cities being sufficient to capture a very large proporation of US traffic remains valid. For more on these issues and the IXmaps project generally, see Clement, A. "IXmaps – Tracking your personal data through the NSA's warrantless wiretapping sites" *IEEE - ISTAS conference*, Toronto, June 26-27, 2013. https://www.dropbox.com/s/9y4xtavova2qtj4/ISTAS13 paper 26 IXmaps %E2%80%93 Tracking May 22.pdf

traceroutes with distinctive icons to highlight those hops most susceptible to NSA splitting. Table 1 and Figure 7 show a traceroute (TR 1859), from a home in Toronto to the San Francisco Art Institute, with hops in the AT&T facilities in both San Francisco and Chicago flagged as likely sites of NSA interception along the way.

**Traceroute detail**

Traceroute id: **1859**
origin: **M5S2M8** destination: **sfai.edu** [63.197.251.33]
submitted by: AndrewC submitted on: 2009-12-13 12:06:51-05

| Hop | IP Address | Min. Latency | Carrier | Location | GeoPrecision | Hostname |
|---|---|---|---|---|---|---|
| 0 | 206.248.154.0 | 0 | TekSavvy | Toronto ON | city level | 206.248.154.0 |
| 1 | 69.196.136.66 | 0 | TekSavvy | Toronto ON | city level | 2120.ac0.bdr02.tor.packetflow.ca |
| 2 | 64.34.236.121 | 0 | Peer 1 | Toronto ON | city level | 64.34.236.121 |
| 3 | 216.187.114.145 | 0 | Peer 1 | Toronto ON | building level | 10ge.xe-2-0-0.tor-151f-cor-1.peer1.net |
| 4 | 216.187.114.133 | 0 | Peer 1 | Toronto ON | building level | 10ge.xe-0-0-0.tor-1yg-cor-1.peer1.net |
| 5 | 216.187.114.141 | 15 | Peer 1 | Chicago IL | building level | oc48-po5-0.chi-eqx-dis-1.peer1.net |
| 6 | 206.223.119.79 | 15 | ESNET - ESnet | Chicago IL | building level | cx1-g1-0.eqchil.sbcglobal.net |
| 7 | 151.164.99.110 | 15 | AT&T Internet Services | Chicago IL | city level | 151.164.99.110 |
| 8 | 151.164.99.129 | 15 | AT&T Internet Services | Chicago IL | city level | 151.164.99.129 |
| 9 | 12.122.79.85 | 15 | AT&T WorldNet Services | Chicago IL | city level | gar3.cgcil.ip.att.net |
| 10 | 12.122.133.218 | 62 | AT&T WorldNet Services | Chicago IL | city level | cr1.cgcil.ip.att.net |
| 11 | 12.122.4.121 | 62 | AT&T WorldNet Services | San Francisco CA | building level | cr1.sffca.ip.att.net |
| 12 | 12.123.15.110 | 62 | AT&T WorldNet Services | San Francisco CA | building level | cr83.sffca.ip.att.net |
| 13 | 12.122.110.113 | 62 | AT&T WorldNet Services | San Francisco CA | building level | gar26.sffca.ip.att.net |
| 14 | 12.91.92.250 | 62 | AT&T WorldNet Services | San Francisco CA | building level | 12.91.92.250 |
| 15 | 63.197.251.33 | 62 | AT&T Internet Services | San Francisco CA | Maxmind | 63.197.251.33 |

**Legend**
- NSA: Known NSA listening facility in the city
- NSA: Suspected NSA listening facility in the city

Table 1: Traceroute details for TR #1859, Toronto to San Francisco Art Institute

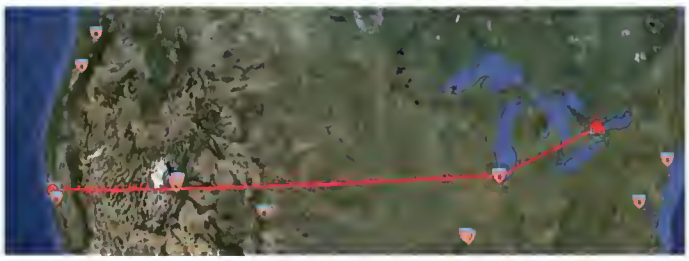

Figure 7: IXmaps rendering of traceroute #1859, Toronto to San Francisco Art Institute

### 4.3 Non-US traffic may also be exposed to NSA splitters

So far we have concentrated on traffic that explicitly travels via US routing centres, i.e. originating or terminating in the US, or both. It is well known, at least in internet circles, that traffic that neither originates nor terminates in the US may nevertheless transit via the US, mainly due to the interconnection arrangements of the major international carriers (Norton, 2012, p. 71). However, the extent of this practice and its surveillance implications are less well known. While this affects many countries, Canadian traffic in particular, largely due to its proximity to the US as well as the structure of the North American internet service industry, is especially prone to routing via the US. We refer to traffic that originates and terminates

in the same country, but transits another, as "boomerang traffic." Analysis of IXmaps data reveals that approximately one third of the Canadian routes follow a boomerang pattern. That long distance Canadian routes may be routed via the US is not surprising, but we were struck by the number of routes that start and end in the same Canadian city, but are routed via the US. We have found over 100 such boomerang routes based in Toronto alone. Figure 8 shows one example that transits New York and Chicago, both cities strongly suspected of hosting NSA splitters. Whether crossing the continent, or returning to the same city, Canadian boomerang traffic is almost entirely exposed to NSA surveillance.

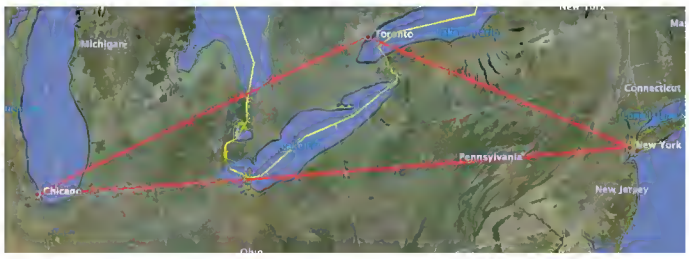


Figure 8: A Canadian boomerang route based in Toronto (TR6896)

## 5 Discussion

When a government conducts surveillance on its citizens and acts outside conventional legal bounds, as the US government has arguably done in the case of the NSA's surveillance programs, the norms of liberal democratic governance are seriously violated and demand public accountability. However, the usual mechanisms for such accountability have not so far succeeded. Congress passed legislation in 2008 that retroactively granted the implicated telecom carriers immunity from prosecution and the executive branch has until recently largely stymied court challenges by invoking a blanket "state secrets" exemption. While over the past decade there have been several notable journalistic exposés (Bamford, 2008; PBS Frontline, 2007) and brave whistleblowers from both the NSA and AT&T have brought damning information to light, it is only months after the Snowden revelations that the public policy debate is getting underway in earnest and may still falter.

This paper has attempted to contribute to this debate by exploring the geographic dimensions of the NSA surveillance programs. To counteract the popular but misleading metaphor of the internet as a 'cloud', we have reviewed the three main NSA interception programs, bulk telephony metadata collection, Prism and Upstream, in each case highlighting the critical role that spatial and locational features play in what data is collected, on whom and how. In aggregate, these interceptions programs are capable of and likely are capturing almost all electronic communication, at least within the U.S.

Compounding the usual difficulties in holding powerful players responsible for their actions is the intrinsically invisible character of internet surveillance. Beyond the notorious secrecy of the NSA, the surveillance is conducted out of sight and leaves no discernible trace. For the great majority of the population, the workings of the internet, especially in its core, are dauntingly complex and inscrutable. We have developed the IXmaps internet mapping application to overcome these obstacles by promoting greater transparency and visibility of the NSA surveillance activities. Within the limits of the available information, we have been able to reveal the likely sites of NSA surveillance operations and show interested individuals where the NSA may intercept their own data packets. By using interactive maps and graphic images, we hope to make the surveillance more vivid, discussable and a matter of public concern.

But more than just serving concerned citizens and curious explorers, IXmaps encourages and relies on its users to contribute to building its database of traceroutes. This crowdsourcing is necessary to ensure a good geographic distribution of originating points, so that the internet core is well surveyed, but also provides the means for people to view the internet from their own personal perspectives. Perhaps more

importantly, integrating these contributions in an open and publicly visible manner constitutes a form of collective counter-surveillance with the potential to empower participants in holding the US government and its national security agency to account for its warrantless surveillance.

We also hope this constructive, surveillance studies approach to the problem of unaccountable state surveillance, can stimulate a critical and productive discussion within the information studies field about its 'darker side' and how we as implicated participants can act responsibly, individually and collectively, in the face of a most serious societal challenge. In particular, widespread faith in the power of harnessing ever more information capturing and processing capabilites to solve societal problems, as exemplified in the NSA programs and current enthusiasm more generally for 'big data' techniques, needs to be tempered by a careful examination of its implications in light of other values, such as privacy and democratic governance.

## 6 References

Bamford, James. (2008). The Shadow Factory: The UltraSecret NSA from 9/11 to the Eavesdropping on America. New York: Doubleday.

Bamford, James. (2012, March 15). "The NSA Is Building the Country's Biggest Spy Center (Watch What You Say)". Wired. http://www.wired.com/threatlevel/2012/03/ff_nsadatacenter/all/1

Dodge. M. and Kitchen, R. (2002). "New Cartographies to Chart Cyberspace" Geoinformatics (April/May):1.

Electronic Frontier Foundation. (2012, July 2). Three NSA Whistleblowers Back EFF's Lawsuit Over Government's Massive Spying Program. https://www.eff.org/press/releases/three-nsa-whistleblowers-back-effs-lawsuit-over-governments-massive-spying-program

Gilliom, J. & Monahan, T. (2013). SuperVision: An Introduction to the Surveillance Society. Chicago: University of Chicago Press.

Katz-Bassett, E. , J. John, A. Krishnamurthy, D. Wetherall, T. Anderson, and Y. Chawathe. (2006, October). Towards IP Geolocation Using Delay and Topology Measurements," in ACM IMC '06.

Klein, Mark. (2009). Wiring up the Big Brother Machine... and fighting it. Charleston, SC: BookSurge.

Landau, Susan. (2011). Surveillance or Security? The Risks Posed by New Wiretapping Technologies, Cambridge MA: MIT Press.

Lyon, David. (2007). Surveillance Studies: An Overview. Cambridge, UK: Polity Press.

Norton, William B. (2012). The Internet Peering Playbook: Connecting to the Core of the Internet, DrPeering Press, http://drpeering.net/core/bookOutline.html.

PBS Frontline. (2007, May 15). Spying on the Home Front. http://www.pbs.org/wgbh/pages/frontline/homefront/view/

Vinson, Judge Roger (2013, April 25). In re Application of the Federal Bureau of Investigation for an Order Requiring the production of tangible things from Verizon Business Network Services, Inc. on behalf of MCI Communication Services, Inc. d/b/a Verizon Business Services (PDF). Top Secret order of the Foreign Intelligence Surveillance Court. Electronic Privacy Information Center. http://epic.org/privacy/nsa/Section-215-Order-to-Verizon.pdf

## 7 Table of Figures

## 8 Table of Tables